JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **224,** 117–130 (1998)
ARTICLE NO. AY985990

Stabilizing Stationary Linear Discrete Systems: Minimal and Balanced Expansions in Any Real Base

Dominique Luzeaux*

Laboratoire Perception pour la robotique, CTA/GIP, 16 bis av. Prieur de la Côte d'Or,

and

Bernard Beauzamy†

Société de Calcul Mathématique, 111 Faubourg St-Honoré, 75008 Paris, France

Submitted by Richard A. Duke

Received September 30, 1996

Let a be a real number strictly greater than 1 and let D be a finite interval of $\mathbb{Z}$ containing 0. In this paper we give necessary and sufficient conditions that guarantee the existence, for any real, of an expansion in base a with coefficients in D. Then we turn to balanced expansions, for which the sums of the digits of any initial segment are uniformly bounded. These issues yield necessary and sufficient conditions for stabilizing stationary linear discrete systems with a particular family of control laws. © 1998 Academic Press

1. INTRODUCTION

This research was motivated by the study of the following dynamical system:

$$\begin{aligned} X_{k+1} &= AX_k + Bu_k, \\ u_{k+1} &= u_k + v_k, \\ v_k &\in \{-m, \dots, +m\}, \end{aligned} \tag{1}$$

*E-mail: luzeaux@etca.fr.
†E-mail: beauzamy@etca.fr.

0022-247X/98 $25.00

where m is an integer. This is a stationary linear discrete system controlled by a rule-based incremental control law [10, 13]. More precisely, X_k is the state of the system, A is the state matrix, and B is the controllability vector; u_k is the scalar input and its variations v_k take only integer values. The problem is to find a sequence of inputs (u_k) such that $\sup_k \|X_k\| < \infty$.

In control theory [19], such an issue is called a state-feedback "bounded-input bounded-state" stabilization of the system $X_{k+1} = AX_k + Bu_k$, with constraints on the inputs given by the recurrent equation satisfied by u_k. Actually, we will deal with controllable systems $X_{k+1} = AX_k + Bu_k$: for every X_0, there exists always a finite sequence of inputs (without constraints on u_k) which drive the system from X_0 to the origin (in other words, there exists k_0 such that if we apply the inputs $u_0, \dots, u_{k_0}$ to the system $X_{k+1} = AX_k + Bu_k$ with initial condition X_0, then $X_{k_0+1} = 0$). One shows that there exists then a coordinate transformation that rewrites the system under a canonical controllable form [19] (A is then a companion matrix and $B = (0 \ \cdots \ 0 \ 1)^{\mathsf{T}}$, where T is the transpose operator). With these choices of A and B, if we take $u_k = -[B^{\mathsf{T}}AX_k]$, where $[\cdot]$ is the integral part, system (1) reduces to $X_{k+1} = \{AX_k\}$. In this new equation, $\{\cdot\}$ stands for the fractional part, and the equality should be understood componentwise.

When A is a scalar, such systems ($X_{k+1} = \{AX_k\}$) are sometimes called β-expansions (following the notation introduced in [16]) and are actually one-sided Bernoulli shifts.[1] Some properties (like the statistical distribution of the orbits) have been studied from an ergodic point of view [4, 7, 15, 16, 17]. These properties yield information on the distributions of the state of the controlled system (1) [12]. In practice such results are interesting: systems which have reached a steady equilibrium state and systems with arbitrarily small state but unpredictable behavior within a small neighborhood are very different! Periodicity properties of the orbits can also be studied from the points of view of topology or measure theory [11], or representability through automata [1, 2, 3]. The case when A is a matrix is covered by the study of higher dimensional Bernoulli shifts, and more generally dynamical systems [5, 9, 14, 18, 20].

In this paper, we define the exact role played by m (the bound on the input variations) in the stability of (1). What we mean by stability will be made precise below. Indeed, we look for the smallest m that guarantees stability for a scalar system (the matrix A is a scalar denoted a from now on). This minimal value of m determines the variations range of the

[1] A *one-sided left Bernoulli shift* transforms a sequence $(q_1, q_2, q_3, \dots)$ into the sequence $(q_2, q_3, \dots)$.

inputs, i.e., the dynamics of the control law: the larger m, the larger the dynamics, and this may be physically hazardous (important dynamics imply commutations at the actuators level, therefore sudden energy dissipations).

First we show that the stability of (1) is equivalent to the existence of what we call a *balanced expansion* of $(a - 1)x_0 + bu_0$, and the core of the paper is then to define such expansions and to relate their existence to a condition between a and m.

More precisely, we have the following theorem:

MAIN THEOREM. *If $m \geq a - 1$, system* (1) *can be stabilized for any value of $a > 1$.*

Conversely, the condition $m \geq [a] - 1$ is necessary for stabilization for all values of a. If a belongs to a particular set of reals (a *union of intervals defined later*), *the condition $m \geq a - 1$ is necessary and sufficient for stabilization.*

Everyone is familiar with the expansion of a number in an integer base (for instance decimal, binary, or hexadecimal expansions): if x is a real and a an integer, then the expansion of x in base a is $x = \sum_{i=-N}^{\infty} d_i a^{-i}$, with $N \in \mathbb{Z}$ and, for all i, $0 \leq d_i < a$ and $d_i \in \mathbb{N}$.

Instead of restricting us to integral values of a, we will consider here any *real* value. Obviously there is no more unicity of the expansion in that case. A first question is then: Which value can be taken by the "digits" d_i? Actually, the question is to know whether these digits have to be arbitrarily large integers. If we refer to decimal expansions, this can be reformulated as: Is it possible to expand a real x in base 10 by using digits taken between two bounds $d_{\min}$ and $d_{\max}$, with $d_{\min} \neq 0$ and $d_{\max} \neq 9$? And what is the minimal difference $d_{\max} - d_{\min}$ for which such an expansion is always possible? Let us mention that we allow negative values for $d_{\min}$ and positive values for $d_{\max}$, since we have to work with any real x, be they positive or negative. This remark is important since, for instance in decimal expansions, we actually use digits between -9 and $+9$ (e.g., $-\pi = -3. -1 -4 -1 -5 \cdots$) and we may wonder whether 19 digits are really necessary to expand any real in base 10 (the answer is *no*).

We will need in the next sections the notion of a *balanced* expansion, i.e., such that the sums of the digits of any initial segment are uniformly bounded: $\sup_{k \geq -N} |\sum_{i=-N}^{k} d_i| < \infty$. For the sake of simplicity, we will assume in that case $d_{\max} = -d_{\min}$.

We will denote by $[x]$ the largest integer that is not greater than x for positive x, and when x is negative $[x] = -[-x]$. Furthermore the difference $x - [x]$ will be denoted by $\{x\}$.

2. FROM STABILITY OF (1) TO EXPANSIONS

From now on, we deal with the scalar case and a controllable system, which is equivalent here to $b \neq 0$. The first state equation of (1) can be rewritten as

$$x_{k+1} = a^{k+1}x_0 + bu_k + abu_{k-1} + \cdots + a^k bu_0.$$

The second equation yields by induction

$$u_{k+1} = u_0 + v_k + v_{k-1} + \cdots + v_0.$$

By combining both these equations, and expressing everything as a function of v_i, we have equivalently

$$\begin{aligned} x_{k+1} &= a^{k+1}x_0 + b(u_0 + v_{k-1} + \cdots + v_0) \\ &\quad + ab(u_0 + v_{k-2} + \cdots + v_0) + \cdots + a^k bu_0 \\ &= a^{k+1}x_0 + bu_0 \frac{a^{k+1} - 1}{a-1} + bv_{k-1}\frac{a-1}{a-1} \\ &\quad + bv_{k-2}\frac{a^2-1}{a-1} + \cdots + bv_0 \frac{a^k - 1}{a-1} \\ &= a^{k+1}\left(x_0 + \frac{bu_0}{a-1}\right) - \frac{bu_0}{a-1} + \frac{b}{a-1} \\ &\quad \times \left(av_{k-1} + a^2 v_{k-2} + \cdots + a^k v_0\right) - \frac{b}{a-1}(v_{k-1} + \cdots + v_0). \end{aligned}$$

After multiplying by $(a-1)$ and dividing by a^{k+1},

$$\begin{aligned} (a-1)x_0 + bu_0 &= \frac{bu_0}{a^{k+1}} + \frac{x_{k+1}(a-1)}{a^{k+1}} \\ &\quad - b\left(\frac{v_0}{a} + \cdots + \frac{v_{k-1}}{a^k}\right) + \frac{b(v_0 + \cdots + v_{k-1})}{a^{k+1}}. \end{aligned} \tag{2}$$

The aim is to stabilize the system (1), more precisely to obtain a BIBS (bounded inputs bounded states [19]) behavior: *we look for a uniformly bounded input sequence such that the state of* (1) *is then uniformly bounded.* From a physical point of view, such a behavior is one of the less constraining one can impose: it is reasonable to consider bounded inputs since they correspond to physical parameters (they drive the effectors of the physical system). As for the state, it should be constrained to domains not too far from the equilibrium (here the origin), which goes with the intuitive meaning of stability (small disturbances only shift slightly the equilibrium).

PROPOSITION 2.1. *System* (1) *will be BIBS stable if and only if the real number* $-((a-1)x_0 + bu_0)/b$ *has, in base* a, *an expansion* $\sum_{k\geq 1} v_{k-1}a^{-k}$ *such that*

$$\sup_{k\geq 1}\left|\sum_{i\geq 1}^{k} v_{i-1}\right| < \infty. \tag{3}$$

Proof. • Let us consider identity (2). Since $v_0 + \cdots + v_{k-1} = u_k - u_0$, this last quantity is bounded if the system is BIBS stable, and (3) is satisfied. Since $bu_0/a^{k+1} \to 0$ and $x_{k+1}(a-1)/a^{k+1} \to 0$, we have indeed $-((a-1)x_0 + bu_0)/b = \sum_{k\geq 1} v_{k-1}a^{-k}$.

• Assume conversely that $-((a-1)x_0 + bu_0)/b = \sum_{k\geq 1} v_{k-1}a^{-k}$, where the v_i satisfy (3). Equation (2) can be rewritten

$$\begin{aligned}
&\frac{x_{k+1}(a-1)}{a^{k+1}} + \frac{b(v_0 + \cdots + v_{k-1})}{a^{k+1}}\\
&\qquad = (a-1)x_0 + bu_0 + b\left(\frac{v_0}{a} + \cdots + \frac{v_{k-1}}{a^k}\right) - \frac{bu_0}{a^{k+1}}\\
&\qquad = -b\left(\frac{v_k}{a^{k+1}} + \frac{v_{k+1}}{a^{k+2}} + \cdots\right) - \frac{bu_0}{a^{k+1}}.
\end{aligned}$$

After multiplication by a^{k+1}, we have

$$x_{k+1}(a-1) + b(v_0 + \cdots + v_{k-1}) = -b\left(v_k + \frac{v_{k+1}}{a} + \cdots\right) - bu_0.$$

Both right terms are bounded; thus if $v_0 + \cdots + v_{k-1}$ is bounded, x_k has to be too. ∎

To sum up, the desired stability of (1) is strictly equivalent to the existence of an expansion $\sum_{k\geq 1} v_{k-1}a^{-k}$ in base a of $-((a-1)x_0 + bu_0)/b$ such that, $\forall k \geq 1$, $|\sum_{i\geq 1}^{k} v_{i-1}| < \infty$. Such expansions, as already mentioned, will be called *balanced*.

3. MINIMAL EXPANSIONS

We assume now that the control v_i can take all integer values between two bounds $d_{\min} < 0$ and $d_{\max} > 0$. We introduce the following notation:

$$A_n = \{d_0 + d_1 a + \cdots + d_n a^n | d_i \in \mathbb{Z},\ d_{\min} \leq d_i \leq d_{\max},\ i = 0, \ldots, n\}$$

is the set of polynomial combinations of a, with a degree smaller than or equal to n and integral coefficients taken between $d_{\min}$ and $d_{\max}$;

$$F = \left\{ x \,\middle|\, x = \sum_{i>0} \frac{d_i}{a^i}, d_i \in \{d_{\min}, \dots, d_{\max}\}, i > 0 \right\}$$

is the set of fractional expansions, with digits taken between $d_{\min}$ and $d_{\max}$.

LEMMA 3.1. *The average distance between two consecutive elements of A_n has the same asymptotic behavior as $(a/(d_{\max} - d_{\min} + 1))^{n+1}$, when $n \to \infty$.*

Proof. The elements of A_n take their values between $d_{\min}(a^{n+1} - 1)/(a-1)$ and $d_{\max}(a^{n+1} - 1)/(a-1)$. The cardinal of A_n, denoted by $\#A_n$, is easy to determine when a is transcendent, since two elements of A_n are then necessarily distinct. Its value is $(d_{\max} - d_{\min} + 1)^{n+1}$.

For algebraic a, there may exist elements of A_n obtained by different choices of $(d_0, \dots, d_n)$, and the value given before is only an upper bound. When this happens, we have an equality such as $d_0^{(1)} + ad_1^{(1)} + \cdots + a^n d_n^{(1)} = d_0^{(2)} + ad_1^{(2)} + \cdots + a^n d_n^{(2)}$. This can be rewritten as $\mathscr{P}(a) = 0$, where $\mathscr{P}$ is a polynomial with unknown a and degree n. Since a is algebraic, there exists a polynomial μ of minimal degree p which admits a as a root. Assume $n > p$. Using the minimal polynomial μ, we can express the powers of a greater than p as a function of the powers of a smaller than p, with integral coefficients. The equality stated before is then equivalent to $\mathscr{Q}(a) = 0$, where $\mathscr{Q}$ has now degree p and still integral coefficients. This may be seen as a set of $p + 1$ linear equations with $n + 1$ unknowns taking integral values between $d_{\min}$ and $d_{\max}$. We conclude that there exist at most $(d_{\max} - d_{\min} + 1)^{n-p}$ choices of d_i that may yield an equality like the previous one. Thus the lower bound is $\#A_n \geq (d_{\max} - d_{\min} + 1)^{n+1} - (d_{\max} - d_{\min} + 1)^{n-p}$.

Since the average distance between two consecutive elements of A_n is the ratio between the difference of the maximal and minimal bounds of A_n and its cardinal, the asymptotic estimate follows. ∎

LEMMA 3.2. *If $d_{\max} - d_{\min} \geq a - 1$, the maximal distance between two consecutive elements of A_n is* 1.

Proof. The proof is made by induction on n.

- By construction, A_0 is the set of all integers between $d_{\min}$ and $d_{\max}$.

- We notice that every element $d_0 + d_1 a + \cdots + d_n a^n$ can be written $(d_0 + d_1 a + \cdots + d_{n-1} a^{n-1}) + d_n a^n$. Therefore

$$A_n = \bigcup_{j \in \{d_{\min}, \dots, d_{\max}\}} (ja^n + A_{n-1}),$$

where the notation $\alpha + B$, for a real α and a set B, represents classically the set of all $\alpha + b$ for $b \in B$. Let us write $A_{n,j} = ja^n + A_{n+1}$.

The various sets $A_{n,j}$ are spread on the real line increasingly with j. The occasional overlapping of $A_{n,j}$ and $A_{n,j+1}$ can be detected by comparing the maximal bound of the former and the minimal bound of the latter:

$$\begin{aligned}
&\inf(A_{n,j+1}) - \sup(A_{n,j}) \\
&\quad = (j+1)a^n + d_{\min}\frac{a^n - 1}{a - 1} - ja^n - d_{\max}\frac{a^n - 1}{a - 1} \\
&\quad = \frac{(-d_{\max} + d_{\min} + a - 1)a^n - d_{\min} + d_{\max}}{a - 1}.
\end{aligned} \tag{4}$$

This expression is equal to 1 for $n = 0$ and decreases then toward $-\infty$ when n increases, since by assumption $d_{\max} - d_{\min} \geq a - 1$; therefore the different sets $A_{n,j}$ overlap. Thus the induction assumption applied to A_{n-1} can be applied directly to the different $A_{n,j} = ja^n + A_{n-1}$ for $j \in \{d_{\min}, \ldots, d_{\max}\}$. The fact that the $A_{n,j}$ overlap allows us to conclude the validity of the induction hypothesis for A_n.

Notice that the maximal distance is reached for all n, as the greatest element (in the usual order meaning) of A_n is $d_{\max} + d_{\max}a + \cdots + d_{\max}a^n$ and the element just before is $(d_{\max} - 1) + d_{\max}a + \cdots + d_{\max}a^n$. ∎

PROPOSITION 3.3. *The condition $d_{\max} - d_{\min} \geq a - 1$ is necessary for the existence of an expansion in base a for every real number.*

Proof. The proof is a reductio ad absurdum. Assume that $d_{\max} - d_{\min} < a - 1$.

In the previous proof, we saw that $A_n = \bigcup_{j \in \{d_{\min}, \ldots, d_{\max}\}} A_{n,j}$. Besides, when $d_{\max} - d_{\min} < a - 1$, Eq. (4) shows that instead of overlapping, both sets $A_{n,j}$ and $A_{n,j+1}$ drift apart following an exponential function of n. Since A_{n-1} is actually $A_{n,0}$, we see that the elements of A_n are those of A_{n-1} as well as other numbers which, on the real line, are strictly on the right or on the left of A_{n-1}. Put in other terms, when going from A_{n-1} to A_n, no new element is inserted between the bounds of A_{n-1}.

If every real x admits an expansion in base a, we have $x = \sum_{i=0}^{N} d_i a^i + \sum_{i>0}^{\infty} d_i/a^i$, thus x is the sum of an element of A_N and an element of F. But $F \subset]d_{\min}/(a-1), d_{\max}/(a-1)[$.

Then let N be such that $\inf(A_N + a^{n+1}) - \sup A_N$ is strictly greater than $d_{\max}/(a-1) - d_{\min}/(a-1)$, which is possible under the assumption $d_{\max} - d_{\min} < a - 1$. All the previous remarks show that there will exist reals x in the interval $]\sup A_N, \inf(A_N + a^{n+1})[$ without expansion. ∎

Remark. If $d_{\min}$ (resp., $d_{\max}$) is taken to be null, the previous lemmata are still valid, and the proposition reformulates as a necessary condition for the existence of an expansion for all positive reals (resp., negative). This follows from the fact that the convex hull of $\bigcup_n A_n$ is no longer equal to the real line in those cases, but is restricted to its positive half (resp., negative).

PROPOSITION 3.4. *The condition $d_{\max} - d_{\min} \geq a - 1$ is sufficient for the existence of an expansion in base a of every real number.*

Proof. Two steps will be needed: first we show that instead of considering any real, we can consider only elements of F; then we describe an algorithm that yields the desired expansion for any element of F.

- Assume $d_{\max} - d_{\min} \geq a - 1$. Lemma 3.2 claims that the maximal distance between two consecutive elements of A_n is 1. Besides, under that same assumption, F's length is at least 1.

It is straightforward that $A_n \supset A_{n-1}$ and that, for all N, there exists n such that $\inf A_n < -N$ and $\sup A_n > N$. Therefore, in order to expand x, we need only find an element a of some A_n for which the distance to x is at most 1. Then $x - a$ will simply be considered as an element of F. As it appears on Fig. 1, the only problem could occur because of the asymmetry of F (because $d_{\min}$ is not necessarily equal to $-d_{\max}$): $x - a$ may not be in

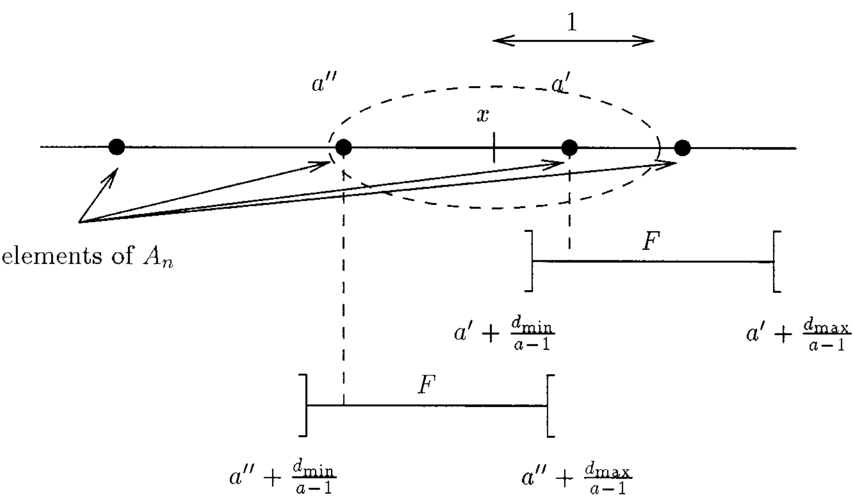

FIG. 1. The real number x is seen as the sum of an element of A_n and an element of F.

F. In order to rule out that unfortunate case, let us introduce $a' = \inf\{a \in A_n | a \geq x\}$ and $a'' = \sup\{a \in A_n | a \leq x\}$. We claim that at least one of the two candidates a' or a'' as an approximation of x is such that the approximating error $x - a'$ or $a'' - x$ is an element of F.

Indeed:

- We know that $a' - a'' \leq 1$.
- If $x - a'' \geq -d_{\min}/(a-1)$, then

$$a' - x - \frac{d_{\max}}{a-1} \leq 1 + a'' - x - \frac{d_{\max}}{a-1} \leq 1 - \frac{d_{\max} - d_{\min}}{a-1} \leq 0.$$

- Similarly, if $a' - x \geq d_{\max}/(a-1)$, then

$$x - a'' + \frac{d_{\min}}{a-1} \leq x + 1 - a' + \frac{d_{\min}}{a-1} \leq 1 - \frac{d_{\max} - d_{\min}}{a-1} \leq 0.$$

- The only thing left is now to give an explicit way of computing the expansion in base a of any element x of $[d_{\min}/(a-1), d_{\max}/(a-1)]$, an interval denoted by I.

Assume we can assign to any z_0 of I a digit $d(z_0)$ taken between $d_{\min}$ and $d_{\max}$, such that $az_0 - d(z_0)$ still belongs to I. If we take then, for all $n \in \mathbb{N}$, $z_{n+1} = az_n - d(z_n)$, we have

$$z_0 = \frac{d(z_0)}{a} + \frac{d(z_1)}{a^2} + \cdots + \frac{d(z_{n-1})}{a^n} + \frac{z_n}{a^n},$$

i.e., an expansion in base a, with a remainder z^n/a^n which behaves like $O(a^{-n-1})$.

Take

$$l = \frac{d_{\max} - d_{\min}}{(a-1)(d_{\max} - d_{\min} + 1)},$$

and, for all $k \in \{0, \ldots, d_{\max} - d_{\min}\}$, define a map d by

$$\begin{cases} \text{if } \dfrac{d_{\min}}{a-1} + kl \leq x < \dfrac{d_{\min}}{a-1} + (k+1)l \qquad \text{then } d(x) = k + d_{\min}, \\ d\left(\dfrac{d_{\max}}{a-1}\right) = d_{\max}, \end{cases}$$

In order to see that such a definition of d yields an expansion in base a, we have to check that for all x in I, we have $ax - d(x)$ in I, i.e.,

$$\forall x \in \left[\frac{d_{\min}}{a-1}, \frac{d_{\max}}{a-1}\right], \qquad \frac{d_{\min}}{a-1} \leq ax - d(x) \leq \frac{d_{\max}}{a-1}.$$

Let us check first the right inequality. We introduce the notation $x \approx y$ to mean that x and y have the same sign:

$$\begin{aligned} &\frac{d_{\max}}{a-1} - ax + d(x) \\ &\qquad \geq \frac{d_{\max}}{a-1} - \frac{ad_{\min}}{a-1} - a\frac{(k+1)(d_{\max} - d_{\min})}{(a-1)(d_{\max} - d_{\min} + 1)} + k + d_{\min} \\ &\qquad \approx (d_{\max} - d_{\min} - k)(d_{\max} - d_{\min} + 1 - a). \end{aligned}$$

But this last expression is always positive, because of the choices of k and a.

Let us check in a similar manner the left inequality.

$$\begin{aligned} ax - d(x) - \frac{d_{\min}}{a-1} &\geq \frac{ad_{\min}}{a-1} + akl - k - d_{\min} - \frac{d_{\min}}{a-1} = k(al-1) \\ &\approx al - 1 \approx d_{\max} - d_{\min} + 1 - a. \end{aligned}$$

This expression is also positive.

We can now conclude the existence of an expansion in base a of all numbers of I with digits between $d_{\min}$ and $d_{\max}$. ∎

Remark. If $d_{\min}$ (resp., $d_{\max}$) is taken to be null, the proposition remains valid under the same restrictions as in the remark following Proposition 3.3.

PROPOSITION 3.5. *The condition $d_{\max} \geq a - 1$ is sufficient for the existence of a balanced expansion in base a for every real number.*

Proof. We need only notice that if $z = \sum_{j \geq -N} d_j/a^j$, the digits being taken between $d'_{\min}$ and $d'_{\max}$, then $z - z/a = \sum_{j \geq -N} (d_j - d_{j-1})/a^j$, with the convention $d_{-N-1} = 0$. Take $d'_j = d_j - d_{j-1}$. It is straightforward that we have an expansion in base a of $z - z/a$ with digits taken between $d'_{\min} - d'_{\max}$ and $d'_{\max} - d'_{\min}$; furthermore the sums of all initial segments are obviously uniformly bounded.

Following Proposition 3.4, as soon as $d'_{\max} - d'_{\min} \geq a - 1$, an expansion of z exists. As $z \mapsto z - z/a$ is one-to-one on $\mathbb{R}$, if we take $d_{\max} = d'_{\max} - d'_{\min}$, we conclude the existence of a balanced expansion for any real in base a with digits taken between $-d_{\max}$ and $d_{\max}$ as soon as $d_{\max} \geq a - 1$. ∎

As a consequence of that last proposition and Proposition 2.1, we have shown the first part of the main theorem, namely the sufficiency of the condition $m \geq a - 1$ for the stabilization of system (1).

4. BALANCED EXPANSIONS

In this section, we deal with the proof of the second part of the theorem. We restrict ourselves to the case $d_{\min} = -d_{\max}$.

PROPOSITION 4.1. *The condition $d_{\max} \geq [a] - 1$ is necessary for the existence of a balanced expansion of any real in base a.*

Proof. Obviously we only need prove the lemma for all positive reals (else invert the sign of all digits in the expansion). The following proof is a reductio ad absurdum.

When $1 < a \leq 2$, the condition $d_{\max} \geq [a] - 1$ reduces in the worst case to $d_{\max} \geq 1$, which is obviously necessary. Then let $a > 2$.

Assume $d_{\max} < [a] - 1$; thus $d_{\max} \leq [a] - 2$. Consider the numbers of the form $\sum_{i=0}^{N} a^i + 1/(a^q(a-1))$ for N and q integers, in other words 1s on the left of the separation mark, followed by a finite number of 0s and an infinite number of 1s. Call d_j the digit at the jth rank (d_j is equal to 0 or 1). Assume that number has another expansion with digits ϵ_j, with $d_{\max} \leq [a] - 2$. Let i be the index of the first digit different in both expansions (by convention, the indices of the digits left of the separation mark will be taken negative, so that the previous numbers are actually written $\sum_{i=-N}^{\infty} d_i/a^i$). Then

$$\left| \sum_{j>i}^{\infty} \frac{\epsilon_j - d_j}{a^j} \right|$$
$$\leq \frac{d_{\max} + 1}{a^{i+1}} \frac{a}{a-1} \leq \frac{[a] - 1}{a^i(a-1)} \leq \frac{1}{a^i} \quad \text{and} \quad \left| \frac{\epsilon_i - d_i}{a^i} \right| \geq \frac{1}{a^i}.$$

We have just given a lower bound for the error introduced by the different digit, and an upper bound for the difference that subsequent digits could eventually have cancelled. Obviously, the only way these bounds join is when all inequalities are equalities. This occurs when a is an integer and all ϵ_j are equal to $2 - a$ from some rank $i + 1$ on, which obviously forbids a uniform bound on the sums of all initial segments. Therefore the condition $d_{\max} \geq [a] - 1$ is necessary. ∎

PROPOSITION 4.2. *For at least all a belonging to the union of intervals defined by the inequality $a < 1/(1 - \{a\})$, the condition $d_{\max} \geq a - 1$ is* *necessary for the existence of a balanced expansion in base a of every real number. Furthermore, the set defined as the union of these intervals has an infinite Lebesgue measure.*

Proof. Once more, the proof is a reductio ad absurdum. Consider the numbers with an expansion in base a of the form

$$0,\underbrace{0\cdots0}_{q_0}1\underbrace{0\cdots0}_{q_1}1\underbrace{0\cdots0}_{q_2}\cdots,$$

i.e., blocks of 0s separated by a 1. As previously, call i the index of the first different digit in both expansions. Then call u_1 the distance to the nearest 1 on the right (the rank of that 1 is thus $i+u_1$), then u_2 the distance to the next 1 (with rank $i+u_2$), and so on. We have

$$\begin{aligned}\left|\sum_{j>i}^{\infty}\frac{\epsilon_j-d_j}{a^j}\right| &\le \sum_{i<j<u_1}\frac{d_{\max}}{a^j}+\frac{d_{\max}+1}{a^{i+u_1}}\\ &\quad+\sum_{i+u_1<j<i+u_2}\frac{d_{\max}}{a^j}+\frac{d_{\max}+1}{a^{i+u_2}}+\cdots\\ &=\sum_{j>i}\frac{d_{\max}}{a^j}+\sum_{p=1}^{\infty}\frac{1}{a^{i+u_p}}\\ &=\frac{[a]-1}{a^i(a-1)}+\frac{1}{a^i}\left(\frac{1}{a^{u_1}}+\sum_{p=2}^{\infty}\frac{1}{a^{u_p}}\right).\end{aligned}$$

It is important to notice that, for $p\ge 2$, the u_p can be chosen arbitrarily large, since they are directly related to the q_l (the choice of which is unconstrained). On the other hand u_1 cannot be chosen randomly, because it corresponds to the difference between both expansions. Therefore $1/a^{u_1}\le 1/a$ and $\sum_{p=2}^{\infty}1/a^{u_p}$, denoted by $S(a)$, can be taken arbitrarily small.

In order to prove that if $d_{\max}=[a]-1$, we cannot always obtain, for all values of a, an expansion of all reals in base a (and necessarily, for these values of a, $d_{\max}\ge[a]\ge a-1$), we need only find some a such that $|\sum_{j>i}^{\infty}(\epsilon_j-d_j)/a^j|<1/a^i$. Following the preceding discussion, let us look for the existence of a such that

$$\frac{[a]-1}{a^i(a-1)}+\frac{1}{a}+S(a)<1.$$

Replacing a with $[a]+\{a\}$, this condition is rewritten as

$$a<\frac{1}{1-\{a\}}-\frac{a(a-1)S(a)}{1-\{a\}}.$$

Since $S(a)$ can be chosen arbitrarily small, we need only study the inequality

$$a < \frac{1}{1 - \{a\}}. \tag{5}$$

Let us fix $\{a\}$. There is obviously a finite number of solutions of (5), separated by a distance of 1. For another value of $\{a\}$ very near but greater, we have solutions very near the previous ones, and possibly new solutions. When $\{a\}$ goes to 1, the bound on a goes to infinity. All this implies that the set of the values of a satisfying (5) is a union of intervals (which we call elementary) of decreasing length and their right extremities are integers. Let us compute the Lebesgue measure of that set. An elementary interval can be written as $]\alpha, n+1[$, where α is a real between n and $n+1$. More precisely, α is such that, for some a, $\alpha = 1/(1 - \{a\})$. Then, from $n \leq \alpha \leq n+1$ one deduces $1 - 1/n \leq \{a\} \leq 1 - 1/(n+1)$. The length of an elementary interval $]\alpha, n+1[$ behaves thus like $O(n^{-1})$. Since the series $\sum n^{-1}$ diverges, we conclude that the set of solutions of (5) has an infinite measure in $\mathbb{R}$. ∎

This concludes the proof of the second part of the theorem.

5. CONCLUSION AND EXTENSIONS

In this paper, we have first reduced the problem of stabilizing a particular linear discrete system to the existence problem of constrained expansions in any real base a. The necessary and sufficient conditions exhibited allow us to solve the initial stabilization problem. The role played by m in Eq. (1) has been fully studied: in the general case, BIBS stabilization is possible if and only if $m \geq a - 1$, and this yields the finesse of the control applied to the considered system.

It seems natural to extend the domain of the values taken by a, for instance to complex values [6, 8, 11]. Unfortunately, this is not straightforward, since we lose the total order of the real set, which is the keystone to all the previous proofs. If we refer to our initial control theoretic problem, we could also consider matricial values of a (and consider a multidimensional system in place of a scalar system). Although it is possible to define the concept of a matricial expansion, the notions of minimality and balance do not translate easily. In the general case, we have found control laws that stabilize the controlled system (which corresponds in the scalar case to finding a balanced expansion) [10]. As a comparison, when these control laws are restricted to the scalar case, they yield the condition

$d_{\max} \geq 2a$: there is a significant loss compared to the case discussed in this paper.

REFERENCES

1. D. Berend and C. Frougny, "Computability by Finite Automata and Pisot Bases," Technical report, LITP, 1992.
2. A. Bertrand-Mathis, Développement en base θ, répartition modulo un de la suite $(x\theta^n)_{n \geq 0}$, langages codés et θ-shift, *Bull. Soc. Math. France* **114** (1986), 271–323.
3. F. Blanchard, β-expansions and symbolic dynamics, *Theoret. Comput. Sci.* **65** (1989).
4. I. P. Cornfeld, S. V. Fomin, and Ya. G. Sinai, "Ergodic theory," Comprehensive Studies in Mathematics, Vol. 245, Springer-Verlag, Berlin/New York, 1982.
5. J. de Vries, Ed., "Elements of Topological Dynamics," Kluwer Academic, Dordrecht/Norwell, MA, 1993.
6. G. A. Edgar, "Measure, Topology and Fractal Geometry," Undergraduate Texts in Mathematics, Springer-Verlag, Berlin/New York, 1990.
7. J. N. Franklin, On the equidistribution of pseudo-random numbers, *Quart. Appl. Math.* **16** (1958).
8. D. E. Knuth, "The Art of Computer Programming," Vol. 2, Addison-Wesley, Reading, MA, 1981.
9. D. Lind and B. Marcus, "Symbolic Dynamics and Coding," Cambridge Univ. Press, Cambridge, UK, 1995.
10. D. Luzeaux, "Miscellaneous on Rule-Based Incremental Control: From Stability to Learnability." Technical Report, UCB-ERL-M92-86, University of California, Berkeley, 1992.
11. D. Luzeaux, From beta-expansions to chaos and fractals, *in* "From Biology to Computation" (D. G. Green and T. J. Bossomaier, Eds.), IOS Press, Amsterdam, 1993.
12. D. Luzeaux, Intelligent control, ergodicity and chaos, *in* "International Simulation Technology Multiconference, Simtech-Wnn-Fnn93," San Francisco, CA, 1993.
13. D. Luzeaux, Process control and machine learning: Rule-based incremental control, *IEEE Trans. Automat. Control* **39**(6) (1994).
14. R. Mané, "Ergodic Theory and Differentiable Dynamics," Series of Modern Mathematics, Vol. 8, Springer-Verlag, Berlin/New York, 1987.
15. W. Parry, On the β-expansion of real numbers, *Acta Math. Acad. Sci. Hungar.* **11** (1960), 401–416.
16. A. Renyi, Representation for real numbers and their ergodic properties, *Acta Math. Acad. Sci. Hungar.* **8** (1957), 477–493.
17. P. Roos, Iterierte Resttransformationen von Zahldarstellungen, *Z. Wahrsch.* **4** (1965), 45–63.
18. Ya. G. Sinaï, Ed., "Dynamical Systems," Vol. II, Encyclopedia of Mathematical Sciences, Springer-Verlag, Berlin/New York, 1989.
19. E. D. Sontag, "Mathematical Control Theory: Deterministic Finite Dimensional Systems," Texts in Applied Mathematics, Vol. 6, Springer-Verlag, Berlin/New York, 1990.
20. Y. Takahashi, An ergodic-theoretical approach to the chaotic behavior of dynamical systems, *Publ. Res. Inst. Math. Sci.* **19** (1983), 1265–1282.